

大小城市合并与行政边界地区经济增长： 基于机器学习算法的合成控制评估

郭峰 吕斌 熊云军 陶旭辉

【摘要】城市行政边界地区在经济发展中处于弱势地位,行政区划调整则为边界地区经济增长带来新的契机,但政策效应在边界两侧可能是不对称的。本文以“莱芜并入济南”作为一次自然实验,利用细颗粒度的卫星灯光数据和基于机器学习算法的合成控制法,评估了大、小两个城市合并对原边界地区经济增长产生的影响。结果显示,济南莱芜合并对原边界地区经济增长产生正向影响,边界地区灯光亮度的实际值比基于机器学习算法预测的反事实合成值平均高出 10.8%。这一结论在更换了边界地区度量方式、对照组选取方式、机器学习算法等后依旧保持稳健。进一步研究还发现,城市合并的正向效应主要体现在核心城市济南市边界一侧,而被合并的小城市莱芜市边界一侧则受益甚微。另外,本文也通过预留样本、考察其他案例的方法验证了核心结论的外部有效性,表明基于机器学习算法的合成控制法有较为广泛的适用性。

【关键词】行政边界;卫星灯光数据;合成控制法;机器学习

【作者简介】郭峰,上海财经大学公共经济与管理学院教授,E-mail:guo.feng@mail.sufe.edu.cn;吕斌,上海财经大学公共经济与管理学院博士研究生,E-mail:lb648371305@foxmail.com;熊云军,西南财经大学经济学院讲师,E-mail:yjxiong@163.sufe.edu.cn;陶旭辉(通讯作者),浙江工商大学公共管理学院讲师,浙江工商大学统计数据工程技术与应用协同创新中心,E-mail:tao_xuhui@126.com。

【原文出处】《数量经济技术经济研究》(京),2024.9.26~48

【基金项目】本文获得上海市哲学社会科学规划课题(2023ZJB006)和浙江省教育厅科研项目(Y202353407)的资助。

一、引言

城市是现代经济体系的骨架和非农产业发展的重要载体,对经济增长具有至关重要的作用(Henderson和Wang,2005;Lucas,2004)。1981-2022年,中国地级市数量由110个增长至293个,城市建成区面积由7438平方公里增加至6.37万平方公里,城镇人口也从2亿增长到9亿^①。在城市数量、面积和人口等各个方面,中国城市建设的规模和速度都位居世界前列。然而,在城市的发展过程中,城市的规模不仅由市场所决定,也深受行政力量的影响,包括城市行政区划的约束(唐为,2019;许政等,2010)。大量文献认为中国的地方政府官员存在围绕经济增长的竞争(Xu,2011;周黎安,2007),行政区划的存在必然会形成行政边界地区,如果在行政边界地区投入更多的资源,也会为相邻地区带来“溢出效应”。因此地方官员往往减少对行政边界地区的资源投放,从而这些行政边界地区的发展通常较为落后,例如上海和苏州边界地区,以及北京和廊坊边界地区。

为突破行政边界对城市规模扩张的约束,加快城市化进程的脚步,近年来,城市行政区划调整频繁进行(唐为和王媛,2015)。城市行政区划调整极大

地拓展了城市发展空间的行政约束,助力城市规模的发展壮大。行政区划的调整不仅包括城市在自己的行政区划内部进行的区划调整优化,例如撤县设区(唐为,2019;唐为和王媛,2015)、开发区设立(魏守华等,2020)、国家级新区设立(郭峰等,2023a)等,也包括城市与城市之间的行政区划调整,例如不同城市之间的直接合并(付婷婷和张同斌,2022;李宝礼和邵帅,2019)。然而,在行政边界地区发展较为落后的现实情况下,城市行政区划的调整,特别是大城市兼并小城市这种情形,究竟能否改变这些地区的落后现状,则是一个需要认真考察的问题。目前,已有部分文献研究了城市内部行政区划调整是否促进边界地区的发展(唐为,2019),而鲜有文献考察城市间的行政区划调整对边界地区经济增长的影响。因此不同城市之间的合并能否促进原边界地区更快发展,是本文关注的重点问题。此外,城市规模直接决定了城市集聚效应的强弱(魏守华等,2015),而大小城市的合并是同时对原边界地区两侧经济增长产生了相同影响,还是对两侧可能存在政策效应的异质性,是另一个亟待研究的关键问题。

本文以山东省省会济南市2018年年末合并地

级市莱芜市为案例,考察了大小城市合并对原城市行政边界地区经济增长的影响。具体来说,本文利用2012—2021年月度夜间卫星灯光数据,考察了济南莱芜合并对他们的原边界地区经济增长的因果关系,同时更为细致地讨论了城市合并对边界两侧的异质性政策效应。对于类似上述仅存在单个独特的处理组情形,由Abadie和Gardeazabal(2003)、Abadie等(2010)提出的合成控制方法是评估这一政策效应的最佳选择。但是,在本文的应用场景中存在三个特殊情况:一是月度卫星灯光数据存在一定的季节性波动,数据不太稳定;二是在城市边界细颗粒度区域层面和月度频率上,缺少相应的协变量;三是处理组和控制组的结果变量可能存在系统性差异。第一种情况在传统合成控制框架下会带来“内插偏差”(Interpolation Bias)的问题,即因预测能力不足而无法很好地复制政策干预前的结果变量;第二种情况在传统合成控制框架下会带来“过拟合”(Overfitting)的问题,这会导致无法判断干预后产生的结果是否是因为过度拟合了控制组;第三种情况的存在可能会导致无法获得最优权重解的问题。前两个特殊情况都会带来“伪政策效果”的风险,最后一个情况则可能会导致传统合成控制方法失效。为此,本文借鉴Doudchenko和Imbens(2016)等文献的思路,基于机器学习算法(弹性网络算法、Lasso算法)构造合成控制,以解决上述问题。研究结果表明,莱芜市并入济南市这一外生政策冲击确实促进了两城市原边界地区经济增长:济南莱芜原边界地区实际增长趋势高于采用机器学习算法合成的反事实路径。具体而言,济南莱芜边界地区灯光亮度实际值比根据机器学习算法预测的反事实合成值平均高了10.8%。但同时,两城市合并对原边界两侧产生了一种非对称的政策效应:城市合并对边界地区经济增长的促进作用主要体现在大城市济南市边界一侧,而对被合并的莱芜市边界一侧经济增长的正向影响并不明显。

本文的研究贡献主要体现在以下三个方面。第一,受限于数据颗粒度的限制,研究行政区划边界地区经济增长的文献考察对象大多数是省际边界地区。与以往文献不同,本文的研究进一步细致到现代经济生态的基本单元——城市层面。使用夜间卫星灯光数据,可以帮助我们对城市边界地区的经济活动进行更细致的刻画,从而得到关于城市合并对边界地区经济增长影响的更直接证据。第二,关于城市行政区划调整,大多数文献主要研究城市辖区内部的区划调整,而鲜有文献研究城市间合并产生的影响。本文的研究为此提供了重要补充。而且,本文还重点探讨了“大小城市合并”对原城市边界两侧产生的异质性影响,为行政边界调整相关文献

提供了重要补充。第三,本文借鉴Doudchenko和Imbens(2016)等文献的做法,构建基于机器学习算法的合成控制法,拓展了合成控制法的适用边界,这也是本文的重要创新。相比传统合成控制方法,基于机器学习算法的合成控制法可以有效缓解因数据波动性大和缺少协变量而导致的“伪政策结果”,以及因处理组和控制组之间存在系统性差异而无法获得权重解的问题。更为重要的是,本文创新性地提出在合成控制方法的应用中将一部分样本预留出来作为模型外测试集,通过观察合成结果和真实结果在测试集上的表现,可以排除传统合成控制方法运用中因为过度拟合而产生的“伪政策效果”问题。

本文余下内容结构安排如下:第二部分是理论分析与政策背景;第三部分重点阐述如何使用机器学习算法更好地进行合成控制,并简要介绍本文使用的主要数据;第四部分是实证结果分析;第五部分是进一步分析,讨论了边界两侧异质性、中心城区异质性以及长春公主岭案例等;第六部分总结了本文的研究结论与相关政策启示。

二、理论分析与政策背景

(一) 理论分析

本文研究主题是城市合并对城市原行政边界地区经济增长的影响,而大量文献已经研究发现行政边界地区的经济增长相对缓慢,经济增长水平和速度普遍落后于非行政边界地区。例如,王文凯和任元明(2022)、唐为(2019)、周黎安和陶婧(2011)等都发现中国的省际边界地区经济增长相对滞后。郭峰等(2023b)发现,行政边界发展落后的现象不仅仅存在于省际边界,同样也存在于城市间的行政边界地区。行政边界地区的发展滞后已经成为制约国内经济大循环和区域协调发展的主要障碍之一,即便是地区发展一体化的口号已经提出多年,行政边界地区始终未得到有效的发展,反而形成了“沿边界贫困带”。

关于边界地区发展缓慢的原因,除了地理障碍方面的因素外,最主要的是来自行政权分割和地方保护主义方面的因素。行政区划的存在势必导致地区间出现多个同级的地方政府,从而增加了地方政府的竞争程度,而边界地区的经济增长不仅有利于当地政府获得更好政绩,也会给属于竞争对手的邻近地区产生溢出效应。因而行政边界地区往往受到地方政府和官员有意或无意的忽视,从而导致边界地区基础设施(尹恒和徐琰超,2011)、金融资源(郭峰和胡军,2016)等投入都相对较少,严重阻碍了行政边界地区的经济发展(唐为,2019;周黎安和陶婧,2011)。而在城市经济增长中,行政边界的存在也阻碍了要素在不同城市之间的高效流动,这往往会导致城市经济的溢出效应或虹吸效应止步于行政边

界。例如,许政等(2010)研究发现,某城市与区域性大城市距离越近,越有利于该城市的经济增长,但省级行政边界的存在降低了区域性大城市对外省城市的溢出效应。郭峰等(2023a)也发现,国家级新区设立后,会吸引周边地区政府将经济活动往新区边界外围地区配置,以便在享受新区政策红利辐射的同时,又能将经济活动留在本辖区。因此,随着中国城市化的快速推进,优化城市行政区划设置就成为推进中国城市规模扩张的一个重要手段。

对于城市的行政区划调整来说,主要包括内部调整和外部扩张两种方式。在内部调整方面,城市行政区划的调整主要包括撤县设区、两区合并和设立开发区等情形,其中又以撤县设区对城市规模扩张的影响最大。然而,学者们对“撤县设区”促进地区经济增长的实际效果仍存在较大争议。例如,唐为和王媛(2015)认为“撤县设区”显著提高了市辖区城镇常住人口的增长率,有利于提升城市的市场潜力和集聚效应,而卢盛峰和陈思霞(2017)的研究则认为“撤县设区”弱化了县域政府的经济事务激励,将使得辖区内企业的融资约束状况相对于其他县市企业变得更严重。相比内部调整,城市的外部扩张对缓解城市规模扩张面临的约束意义更大,从而有可能对城市经济增长产生更大影响。Ahlfeldt等(2015)利用柏林墙建立和拆除的自然实验,发现柏林墙拆除会促进两地区的经济密度和房价上升。李宝礼和邵帅(2019)发现巢湖市并入合肥市扩大了合肥市的空間范围,从而促进合肥市的经济增长,但付婷婷和张同斌(2022)基于更多省会对外扩张样本的研究却发现,城市对外扩张对长期的经济增长而言是抑制的作用。此外,也有文献考察城市行政合并对地方财政支出等的影响(Blesse和Baskaran, 2016;Reingewertz, 2012),认为城市行政合并有助于减少财政开支。

考虑到行政区划对城市经济增长的重要影响以及边界效应的广泛存在,行政区划的调整合并对原边界地区的经济增长必然产生重要影响,但此方面的文献却相对较少。唐为(2019)利用夜间灯光数据,考察了“撤县设区”政策对行政边界地区经济活动的影响,研究发现“撤县设区”显著提高了原市辖区与被撤并县边界地区的经济活动水平,即降低了区县的边界效应。此外,余华义等(2021)基于上海市静安区和闸北区合并的研究也发现,两区合并在一定程度上稀释了边界效应,导致这些原本属于房价洼地的地区房价提升。张楠等(2023)也发现突破城市内部的行政区划有助于提升本辖区的经济活力,并且对相邻辖区存在溢出效应。然而,尽管以往研究文献考察了城市内部行政区划调整合并对辖区边界地区所带来的一些经济影响,但由于城市间合

并的案例相对较少,因而鲜有文献考察城市间的行政合并是否促进了两城市原边界地区的经济增长。因此,不同城市的行政合并是否促进了原边界地区的经济增长,仍需进行严谨的实证考察。本文以济南合并莱芜这一特殊研究场景进行研究。原济南市和原莱芜市城市规模差异很大,是一种典型的大小城市合并案例,其是否可能对边界两侧产生不对称的影响,也是本文将要考察的一个重点内容。

(二)政策背景

本文以济南和莱芜合并来考察大小城市合并对其原边界地区经济增长产生的影响。济南作为山东省省会,是黄河中下游的中心城市,同时也是全国15个副省级城市之一。截至2018年底,即莱芜并入济南前夕,济南市下辖8区2县,面积为7998平方公里,常住人口为746.04万人,地区生产总值为7856.56亿元,占山东省全省的比重为10.27%。对比国内其他省会城市,济南市的面积和人口规模一直比较靠后,而且还存在首位度低、综合竞争力不强、带动力不足的特点。因此,为了实现资源优化配置,近年来在全国做大省会城市的大背景下(付婷婷和张同斌, 2022),山东省提出要把省会建设摆在突出位置,举全省之力推进“省会战略”,并将其写入山东省政府工作报告。而通过对外扩张,摆脱行政区划对济南市做大做强的限制,就变得非常重要。

面积仅2246平方公里的原莱芜市则是山东省面积最小的地级市,建市之初为泰安市所代管的县级市,1992年才升格为地级市。截至2018年底,莱芜市仅下辖2区,常住人口为137.90万人,地区生产总值为1005.65亿元,其常住人口和地区生产总值均居山东省17个地级市的末位。原莱芜市由于体量小,产业结构单一,对钢铁产业存在过度依赖,在人才吸引力、发展空间方面均处于劣势地位。但是,莱芜市也存在靠近济南市的优良条件,因而成为济南市向外扩张的首选方向。2018年12月26日,国务院正式批复同意济南市和莱芜市的行政区划调整方案:撤销莱芜市,并将莱芜市所下辖区划归济南市管辖。调整后济南市下辖10区2县,地区生产总值到2021年已达到1.14万亿元,在全国省会城市中排名第8位。

济南和莱芜的合并,为做大济南省会的城市规模提供了直接助力,但这种城市合并,是否影响以及如何影响它们之间原来的边界地区的经济增长,则仍有待进一步探究。为了更精确地识别莱芜并入济南对原边界地区经济增长的因果效应,本文挑选了与莱芜和济南地理特征、经济增长相似的山东省内其他城市与城市之间的边界地区作为控制组。具体来说,在本文基准模型中,处理组是济南莱芜边界镇所辖地区^②,而控制组是含原莱芜市在内的17个山

东省地级市产生的所有接壤的两个城市之间的边界镇地区,如济南—淄博边界、济南—德州边界、青岛—烟台边界等,共计 31 个城市边界地区。这 31 个城市边界地区将被用于合成如果没有济南莱芜的合并,济南莱芜原边界地区将具有的反事实增长路径。

三、模型设定与数据说明

(一) 模型设定

合成控制方法被认为是过去 15 年来政策评估文献中最重要的创新 (Athey 和 Imbens, 2017), 在研究上述问题的相关文献中得到广泛应用 (付婷婷和张同斌, 2022; 李宝礼和邵帅, 2019; 王贤彬和聂海峰, 2010)。合成控制法是由 Abadie 等人提出的一种用于评估“单个处理对象”的政策评估方法 (Abadie 和 Gardeazabal, 2003; Abadie 等, 2010)。其核心思想是,采用数据驱动的方式,对不同控制组单元赋予不同的权重,通过线性模型构造出一个与处理组相似的合成控制组,将控制组单元合成的结果在政策干预后的数值看作处理组如果没有被干预的反事实结果,处理组和合成控制组在干预后的差值即为干预政策的政策效果。

由于本文的研究场景存在一些特殊情境,传统合成控制法应用失效。为继续针对该案例进行分析,本文借鉴机器学习算法以拓展合成控制法的适用边界。Doudchenko 和 Imbens (2016) 提出了基于机器学习方法构造合成控制的一般性框架,在他们的设计中,传统合成控制方法下的一些权重、协变量等约束都可以放松,而且在合成时还可以考虑加入截距项,从而大大拓宽合成控制方法的适用场景,同时还能获得更精确的政策发生前合成路径 (Clay 等, 2020)。

具体而言,在本文的研究框架下,合成控制法原理可以概况如下:假设能观测到 $N+1$ 个城市对边界地区 T 期的灯光数据 $Y_{i,s}$, 其中第一个边界地区单元 $i=1$ 作为处理组,即济南莱芜边界地区,其余 N 个边界地区单元 $i=2, \dots, N+1$ 作为控制组,即山东省内其余城市与城市边界地区。实际政策冲击发生在 T_0 时期, $1 \leq T_0 < T$, 其中控制组单元在任何时期都不接受处理 (城市合并), 处理组单元在 $s=1, \dots, T_0$ 时期不接受处理,在 $s=T_0+1, \dots, T$ 时期接受处理。式 (1) 展示了分析时所使用的数据结构,其中 $Y_{i,s}^{obs}$ 的含义如式 (2) 所示, $W_{i,s}$ 代表边界地区 i 在 s 期是否接受处理, $Y_{i,s}^1$ 表示城市对边界地区 i 在 s 时期接受政策干预时的潜在结果, $Y_{i,s}^0$ 表示城市对边界地区 i 在 s 时期没有受到政策干预时的潜在结果。 Y 下角标 t 代表处理组单元, c 代表控制组单元, pre 代表在政策干预前的时点, $post$ 代表在政策干预后的时点。

$$Y^{obs} = \begin{pmatrix} Y_{t,pre}^{obs} & Y_{t,post}^{obs} \\ Y_{c,pre}^{obs} & Y_{c,post}^{obs} \end{pmatrix} = \begin{pmatrix} Y_{t,pre}^0 & Y_{t,post}^1 \\ Y_{c,pre}^0 & Y_{c,post}^0 \end{pmatrix} \quad (1)$$

$$Y_{i,s}^{obs} = Y_{i,s}^{w_{i,s}} = \begin{cases} Y_{i,s}^0 & \text{if } W_{i,s} = 0 \\ Y_{i,s}^1 & \text{if } W_{i,s} = 1 \end{cases} \quad (2)$$

在本研究中,济南莱芜边界地区受到政策干预,因而在 $s > T_0$ 期可以观测到潜在结果 $Y_{t,post}^1$, 但是无法观测到如果它没受到政策干预时的潜在结果 $Y_{t,post}^0$, 因而政策评价的关键是如何估计出济南莱芜边界地区在 T_0 期后的反事实结果 $Y_{t,post}^0$ 。在反事实结果被估计出后,可利用观测到的结果 $Y_{t,post}^1$ 减去反事实结果 $Y_{t,post}^0$, 计算出济南莱芜边界地区受政策干预的处理效应为 $\tau_{t,post}$, 如式 (3) 所示。

$$\tau_{t,post} = Y_{t,post}^1 - Y_{t,post}^0 = Y_{t,s}^1 - Y_{t,s}^0 \quad (3)$$

为了估计 $Y_{t,post}^0$, 我们使用 $Y_{t,pre}^0$ 、 $Y_{c,pre}^0$ 、 $Y_{c,post}^0$ 进行计算, 拟合过程如式 (4) 所示, 即用控制单元按照权重 ω_i 与截距项 μ 的线性组合来拟合 $Y_{t,s}^0$ (Clay 等, 2020)。

$$\hat{Y}_{t,s}^0 = \mu + \sum_{i=2}^N \omega_i \cdot Y_{i,s}^{obs} \quad (4)$$

在估计反事实结果过程中,传统合成控制法是以线性函数形式进行合成,主要通过回归 $Y_{t,pre}^0$ 和 $Y_{c,pre}^0$ 来估计 μ 和 ω , 具体如式 (5) 所示。而且,为了获得权重解,传统合成控制法一般都还对合成系数有一定要求,例如没有截距项、权重系数和为 1、权重非负,以及进一步包含协变量,要求协变量和结果变量满足相同的权重约束等。

$$[\hat{\mu}(\alpha, \lambda), \hat{\omega}(\alpha, \lambda)] = \underset{\mu, \omega}{\operatorname{argmin}} \{ (Y_{t,pre}^{obs} - \mu - \omega' Y_{c,pre}^{obs})^2 \} \quad (5)$$

但是,在本文的情境中,有三个特殊情形,导致使用传统的合成控制法存在很大挑战。首先,本文采用的是月度灯光数据,受到很强的诸如气候等季节性因素的影响,数据很不稳定。这对传统线性合成控制预测能力提出了很高的要求,导致研究人员无法判定真正的政策干预效果 (Athey 和 Imbens, 2019)。合成控制估计量是需要渐近无偏的,但如果合成控制所估计的权重无法完全地复制处理组干预前的结果变量,则会带来“内插偏差”的问题 (Abadie 等, 2010)。其次,由于采用了月度数据以及镇层级的数据,缺少特征协变量。可能带来传统方法失效以及“伪政策效果”的风险,即我们没办法判断估计得到的政策结果是因为上述偏差带来的,还是政策本身的结果,这也是经典的“偏差—方差”权衡问题 (Mühlbach 和 Nielsen, 2021)。最后,根据数据描述,可以发现本文所关注的济南莱芜边界地区与山东其他边界地区的经济发展程度存在系统性的差异,即济南莱芜灯光亮度显著更低。当合成控制法中潜在的控制组个体数量较少且处理组个体属于异常点 (与控

制组存在系统差异)时,传统合成控制法由于有着大量的约束条件,可能会出现权重参数无解的情况,也即无法合成的情境。比如中国上海,其特殊的经济特征使得传统的合成控制方法无法通过其他城市的房价加权合成上海的房价(刘甲炎和范子英,2013)。

因此,本文主要借鉴 Doudchenko 和 Imbens (2016)等文献提出的基于机器学习进行合成控制的方法。这种方法允许对合成参数施加更小的限制,例如合成过程中无须施加权重必须非负的假定,其原因是处理组和控制组的相关性为正,但不代表其偏相关为正,从而放松非负权重的假定是合理的。除此之外,对权重和为1以及不能有截距项等假定的放松,也使得基于机器学习的合成控制法可以得到更灵活且更精确的政策发生前合成路径(Clay等,2020)。本质上,上述方式是对传统合成控制法的一种正则化扩展形式,通过添加正则化项改变目标函数的方式,实现提高模型泛化性能的目的(郭峰和陶旭辉,2023)。此外,为提升模型的预测能力,机器学习方法还会将数据划分为训练集和测试集。训练集用来计算模型参数,测试集则用来评价机器学习模型性能。这些特殊的操作,一方面可以解决由于缺少协变量可能导致的过度拟合问题;另一方面还提升了机器学习预测能力,从而可以更好地合成干预前期具有较大波动的结果变量。而且,机器学习的数据划分方式提示我们可以将政策干预前的样本留出一部分作为测试集。这种方式,可以一定程度识别在干预期后处理组和合成控制组产生的差异,究竟是训练集上过度拟合因素所导致的,还是由于真实政策影响所导致的,降低“伪政策效果”的风险。另外,对于处理组是一个异常点,或者它和控制单元存在系统性差异导致权重参数无解时,Doudchenko 和 Imbens (2016)提出的机器学习下的合成控制方法可以通过添加截距项、允许权重值为负数、权重和不大于1三种放松约束的形式来解决,即允许处理组和控制组存在系统差异,并获得负数权重解。

简而言之,基于机器学习算法的合成控制法允许我们对参数施加更少的约束,取而代之的是增加了权重参数的正则化要求。其中,以弹性网络算法最为典型^③,因此我们这里主要以弹性网络算法为例,讨论基于机器学习算法的合成控制法的基本逻辑和流程。式(6)中展示了利用弹性网络算法进行合成时计算参数的基本思路。

$$[\hat{\mu}(\alpha, \lambda), \hat{\omega}(\alpha, \lambda)] = \underset{\mu, \omega}{\operatorname{argmin}} \left\{ (Y_{t, \text{pre}}^{\text{obs}} - \mu - \omega' Y_{c, \text{pre}}^{\text{obs}})^2 + \lambda \left(\frac{1-\alpha}{2} \|\omega\|_2^2 + \alpha \|\omega\|_1 \right) \right\} \quad (6)$$

其中, α 和 λ 是弹性网络模型中的待调超参数,反映了对合成权重的一种“惩罚”, μ 是待估计的截距

项, ω 是待估计边界地区的权重或合成系数, $\|\omega\|_p$ 代表 ω 的 p 阶正则项。

合成控制的目标是要计算出来 $\hat{\mu}$ 和 $\hat{\omega}$,使得在政策实际发生前,处理组的实际观测结果和合成控制后的结果尽可能相似。在利用式(6)成功计算出 $\hat{\mu}$ 和 $\hat{\omega}$ 后,即可计算控制组个体在处理后的合成结果,也即处理组在政策发生后的反事实结果,如式(7)所示。

$$\hat{Y}_{1,s}^0 = \hat{\mu}(\alpha, \lambda) + \sum_{i=2}^N \hat{\omega}(\alpha, \lambda) \times Y_{i,s}^{\text{obs}} \quad (7)$$

随后一个关键问题是如何选定弹性网络模型的超参数 α 和 λ 。参考 Doudchenko 和 Imbens (2016)、Kumar 和 Liang (2019)、Clay 等 (2020) 等文献的做法,在确定超参数的过程中,依次将除处理组外的每个单元作为伪处理单元(Pseudo-treated Unit),用除处理组单元和该伪处理单元外的其他单元作为控制单元进行合成。在设定机器学习损失函数过程中,将逐个伪处理单元 j 的实际观测值与合成值的差值作为个体损失,并将这些逐个伪处理单元的个体损失值的求和项作为总损失。随后通过代入不同的超参数,观察哪一组超参数能使得总损失函数最小,那么这组参数便是超参数 α 和 λ 的最终取值。具体如式(8)和式(9)所示。此外,为了使得结果更为稳健,在每次合成时还要将样本进行随机 k 等分,每次拿出其中一份作为验证集,其余 $k-1$ 份作为训练集进行训练模型,该过程将重复 k 次,并对 k 次的结果进行平均化处理,从而得到这一次合成的损失函数,该过程也被称为 k 折交叉验证。

$$CV(\alpha, \lambda) = \frac{1}{N} \sum_{j=2}^N [Y_{i,T}^{\text{obs}} - \hat{\mu}(j; \alpha, \lambda) -$$

$$\sum_{i \neq 1, j} \hat{\omega}_i(j; \alpha, \lambda) \cdot Y_{i,T}^{\text{obs}}]^2 \quad (8)$$

$$(\alpha_{\text{opt}}, \lambda_{\text{opt}}) = \underset{\alpha, \lambda}{\operatorname{argmin}} \{ CV(\alpha, \lambda) \} \quad (9)$$

基于机器学习算法的合成控制法,已经在相关实证研究中有所应用。例如, Cole 等 (2020) 在研究武汉封城对当地空气污染水平的影响中,就使用了基于岭回归的合成控制方法,用于在干预前控制组无法很好地合成处理组的情形,并运用结果模型估计由于干预前合成不佳产生的偏差,然后纠正传统合成控制的偏差问题。Guo 和 Zhang (2019) 在研究襄樊改名为襄阳的研究中运用了机器学习方法(LASSO 和弹性网),结果发现更名改革后,襄阳实际GDP年增长率约为1.43%。Mühlbach 和 Nielsen (2021) 提出了结合树模型的合成控制方法,用以分析美国驻以色列大使馆迁址的后果,机器学习在这方面的优势在于提供了当处理组和控制组的结果变量存在未知的非线性特征关系时的解决方案。Kumar 和 Liang (2019) 在研究信贷约束对经济增长的影响时,还利用弹性网络算法来挑选控制组构造

合成控制法,以缓解因干预前期数太少带来的“内插偏差”问题。不仅如此,在机器学习正则化的优势下,合成控制还可以结合双重差分法来应对需要合成的对象“非单一”的场景(Arkhangelsky等,2021)。

(二)数据来源及样本选择

本文研究的是城市合并对原城市行政边界地区经济增长的影响。由于缺乏更细微的数据,传统的指标例如GDP、人口规模等很难区分为边界地区与非边界地区,因而难以使用。而夜间灯光数据则没有这方面的困扰,可以更精确地刻画细小尺度空间单元的经济活动,因此也成为研究行政边界地区经济增长的首选(Jia等,2021;郭峰等,2023b)。关于夜间灯光数据,目前被广泛使用的数据包括两种:DMSP/OLS数据和NPP/VIIRS数据(Gibson等,2020,2021)。与DMSP/OLS数据相比,NPP/VIIRS数据有多个优势:首先,DMSP/OLS数据在2014年2月就停止更新,而NPP/VIIRS数据从2012年4月开始发布至今,而且是月度数据。而莱芜市并入济南市发生于2019年1月,这也就决定了NPP/VIIRS数据作为本文的原始数据更合理。其次,由于DMSP/OLS数据以灰度值来表示夜间平均灯光亮度,是一个离散数据,因而DMSP/OLS数据存在明显的数据“过曝”问题,而NPP/VIIRS数据则是连续的像素辐射值,准确地记录了夜间灯光辐射强度,不存在数据“过曝”问题。最后,NPP/VIIRS卫星过境时间为凌晨1:30,可以有效规避暂时性光源,更好地度量城市建设水平,更适合用来刻画经济活动的空间分布。需要说明的是,受到北半球极昼的影响,NPP/VIIRS数据在每年的6月和7月存在失真的情况,因此本文剔除了每年6月和7月的数据。

在确定使用夜间灯光数据来衡量城市行政边界地区的经济增长后,一个关键问题是,如何科学地确定城市行政边界地区的具体范围。中国的城市行政

区划大小不一,中心城区的位置、规模也不尽相同,因此,主观确定的绝对距离来衡量边界地区的范围很容易带来测量上的误差。更科学的方法是根据城市面积和人口密度动态调整的尺度来确定城市边界地区,而中国的基层行政区划就满足这样的要求。因此,参考郭峰等(2023b)、Jia等(2021)等文献的做法,本文在乡镇层级上来刻画城市的行政边界地区。我们将那些在两个城市行政边界两侧的乡镇定义为边界镇,而这些边界镇所在的辖区就是本文所考察的城市对的边界地区。

(三)数据描述性统计分析

根据山东省境内城市边界地区的灯光亮度分布发现,山东省内城市边界地区的灯光亮度平均值为1.01,最大值为4.85(2014年12月青岛市—日照市),最小值为0.04(2015年5月德州市—滨州市),山东省境内城市边界地区的灯光亮度有着显著差异。

根据图1,我们了解济南莱芜边界地区和其他边界地区灯光亮度(简单平均)随年分布的变化情况。从中可以看出,济南莱芜边界地区月均灯光亮度0.67,最小值为0.21(2015年5月),最大值为1.23(2021年12月),济南莱芜边界地区灯光亮度随着时间推移呈现逐年上升。但总体而言,济南莱芜边界地区灯光亮度明显低于其他边界地区的灯光亮度均值。除此之外,虽然变化趋势比较微弱,但仍可以看出,2019年1月济南莱芜合并之后^④,济南莱芜边界地区灯光亮度与其他边界地区灯光亮度相比,在一定程度上出现了靠拢的趋势,这表明莱芜并入济南这一事件冲击对济南莱芜边界地区经济增长可能存在正向效应,当然这还需要更严谨的实证分析才能得到可信的结论。

四、实证结果分析

(一)基准结果

在基准模型中,我们主要利用机器学习中的弹性网络算法替代传统方法进行合成控制。使用机器

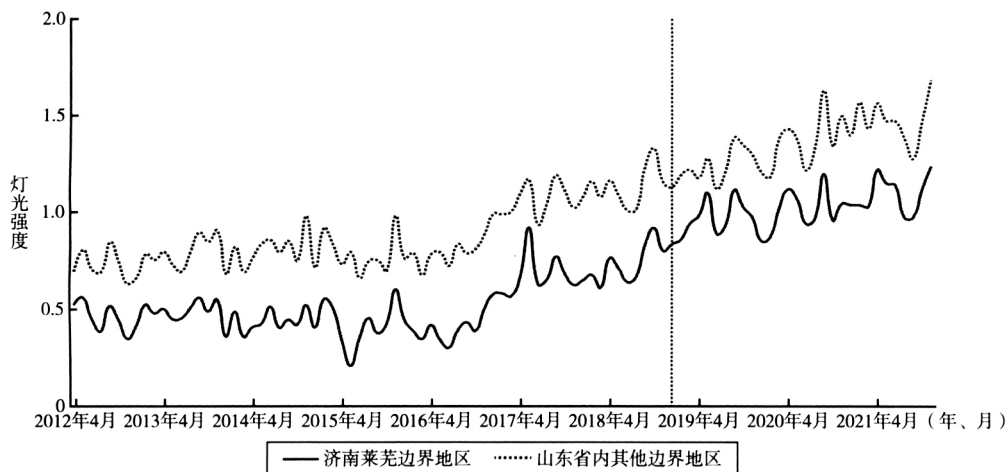


图1 济南莱芜边界地区与其他边界地区的灯光亮度随年分布变化图

学习方法进行合成控制法估计,首先要确定超参数 (Clay 等, 2020; Doudchenko 和 Imbens, 2016; Kumar 和 Liang, 2019; Mühlbach 和 Nielsen, 2021)。按照第三节介绍的方法和步骤,在确定弹性网络模型的超参数过程中,我们依次将山东省内除济南—莱芜(处理组)外的其他城市边界地区作为伪处理单元,并用除处理组以及该伪处理单元外的其他边界地区作为控制组进行合成控制,其间选取全部年份(2012年4月~2021年12月)数据。在完成拟合后,将该伪处理单元的实际观测值与对应控制组单元的合成值相减计算损失。而且,为了增强计算结果的稳定性,在每次合成时,将样本进行五等分,每次拿出其中一份作为测试集,其余作为训练集训练模型,该过程将重复5次,并对5次的结果进行平均化处理,从而得到这一次合成的损失函数。随后把全部伪处理单元的损失相加得到总损失,通过调整超参数来最小化总损失以得到最优参数,该过程即为基准模型的超参数调优过程。

基准模型选取2012年4月~2018年12月的数据作为训练集,并将弹性网络模型在训练集上进行训练合成,随后在2019年1月~2021年12月,使用训练好的机器学习模型和控制组单元数据进行预测,预测的数值结果即为济南莱芜边界地区在处理期后如果没有济南莱芜的合并将具有的反事实结果。结果如图2所示,济南莱芜边界地区的实际观测值(实线部分)与反事实合成值(点线部分)在2019年1月政策实施前较为接近,说明机器学习合成控制法较好地合成了处理组的趋势,而在2019年1月政策实施后,济南莱芜边界地区实际观测结果与合成结果(处理组在政策实施后的反事实结果)发生了一定偏离,而这部分偏离(实线与点线之间的差距),即济南莱芜合并对其原边界地区经济增长的政策效应。图2的结果表明济南和莱芜合并的确促进

了济南莱芜原边界地区的经济增长,并且在量化政策影响时,可发现政策实施后,济南莱芜边界地区灯光亮度的实际值较合成值平均高了10.8%^⑤。根据Henderson等(2012)估计的灯光亮度对收入的弹性,灯光强度增加1%大约相当于GDP增加0.3%。估计结果推广表明,济南和莱芜合并后带来GDP的增长幅度为3.24% ($10.8 \times 0.3 = 3.24$)。考虑到中国的GDP在2012~2021年保持着7.8%的平均增长率,济南和莱芜合并后对于行政边界地区的经济增长影响是巨大的。

正如在前文中所阐述的,由于行政分割等原因,行政边界地区经济发展落后于其他地区,而行政合并消除了或者说至少在某种程度上缓解了这种边界效应背后的行政分割机制。例如,一方面,行政合并之后,原边界地区的正向溢出效应就不复存在,至少对上一级政府而言,无需考虑其在边界地区的投入给竞争对手“搭便车”的机会。因此相对而言,会有更多资源投入到原来的边界“洼地”,实现边界地区相对更快的发展。另一方面,行政合并后,边界地区也从原来地区间的屏障变成合作的前沿阵地,从而能吸引更多企业和其他经济主体入驻,改变原来的发展劣势。总之,原来是发展洼地的行政边界地区,在城市行政合并后会迎来了一个新的发展机遇,实现相对更快的增长,图2的基准模型结果就是上述机制的体现。当然,当两个合并的城市之间规模悬殊较大时,其边界两侧的发展态势是完全不同的,而这种大小城市的合并是否对边界两侧产生异质性的影响,我们将在下文中进行详细阐述。

附表1^⑥中展示了济南莱芜边界地区对各个控制组的合成系数的具体数值,控制组最大系数来自济南—淄博边界地区,系数为0.213。正向系数较大的其他几个城市对边界地区包括淄博—临沂边界地区、临沂—潍坊边界地区、泰安—淄博边界地区、德州—

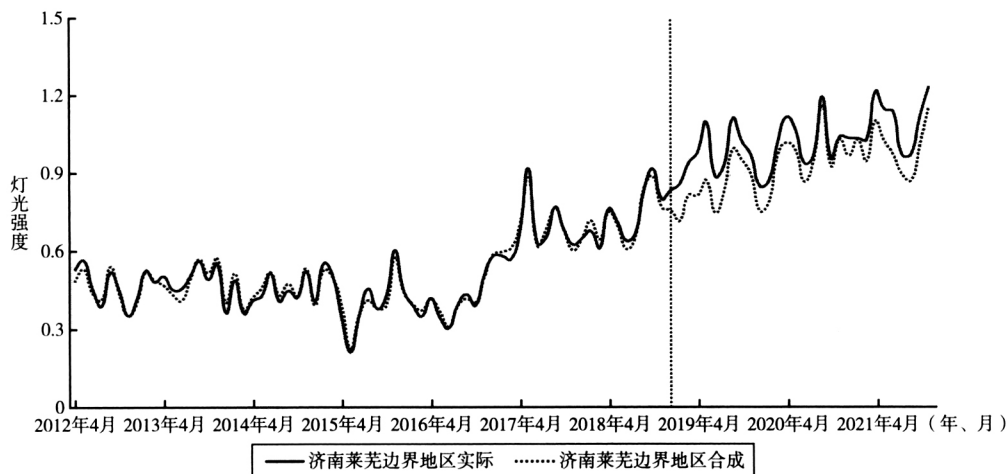


图2 基准模型结果

济南边界地区,而负向系数较大的城市对边界地区包括青岛—烟台边界地区、枣庄—临沂边界地区、临沂—日照边界地区、济宁—枣庄边界地区。共有 24 个城市对边界地区合成系数不为 0。而在弹性网络算法的合成控制基准模型中,截距项系数为-0.074,说明在合成的过程中,处理组和对照组之间确实存在系统性差异。

(二) 稳健性分析

在这个部分,我们通过更换城市边界地区度量方式、更换对照组样本以及采用其他机器学习算法等来检验本文实证结论的稳健性。首先,在边界地区度量方面,本文使用边界镇所辖地区来界定行政边界地区,除这种方法外,文献中也常用绝对距离法来定义行政边界地区。参考刘修岩等(2016)等文献的做法,我们将距离城市边界内 3 千米范围的区域作为行政边界地区,然后重新进行合成控制。此时的实证结果如图 3 所示,从中可以看出政策发生后,处理组实际趋势相对于合成组的反事实趋势平均增

长了 19.98%。虽然使用不同的边界地区度量办法得到的政策效果在绝对数值上存在一些差异,但总体而言采用不同边界地区度量方法得到的结论仍然是稳健的。

其次,在对照组选取方面,我们为济南莱芜边界地区分配的对照组是山东省其他两两城市对的接壤地区,这些地区与济南莱芜边界地区经济发展水平、经济地理特征等更为相似,但不足之处在于对照组数量较少(31 个),对于依赖数据驱动的机器学习算法,相对不足。因此我们也将参与合成控制的区域由山东省内两两接壤地区扩展到全国两两城市边界地区,此时共计 959 个城市边界地区。此时的结果如图 4 所示,从中可以看出政策发生后,济南莱芜边界地区实际结果相对于反事实结果上升了 18.25%,说明在更换了控制组情况下,本文结果依然稳健。这一结果其实也说明,对于基于机器学习算法的合成控制法,由于其数据驱动的特征,对照组的选取并不会引起结果的根本性变化。

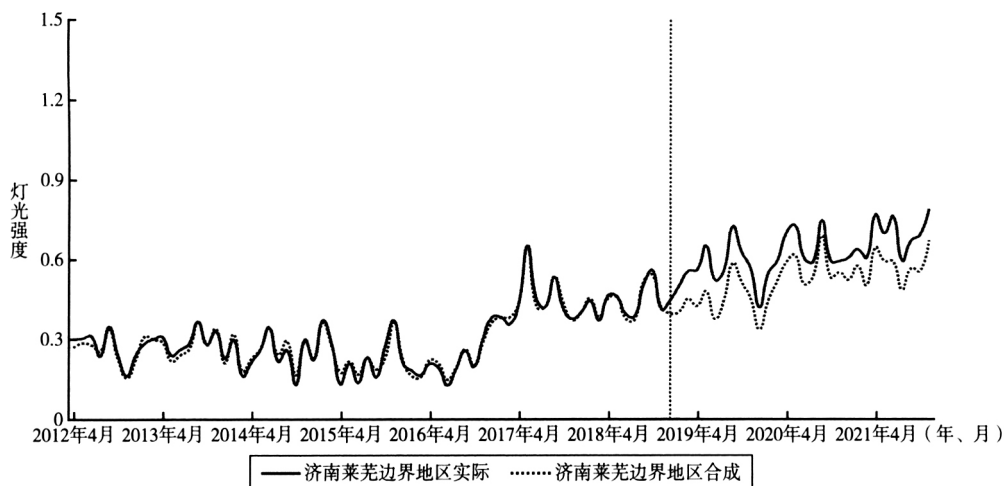


图 3 更换边界地区度量方式:绝对距离法

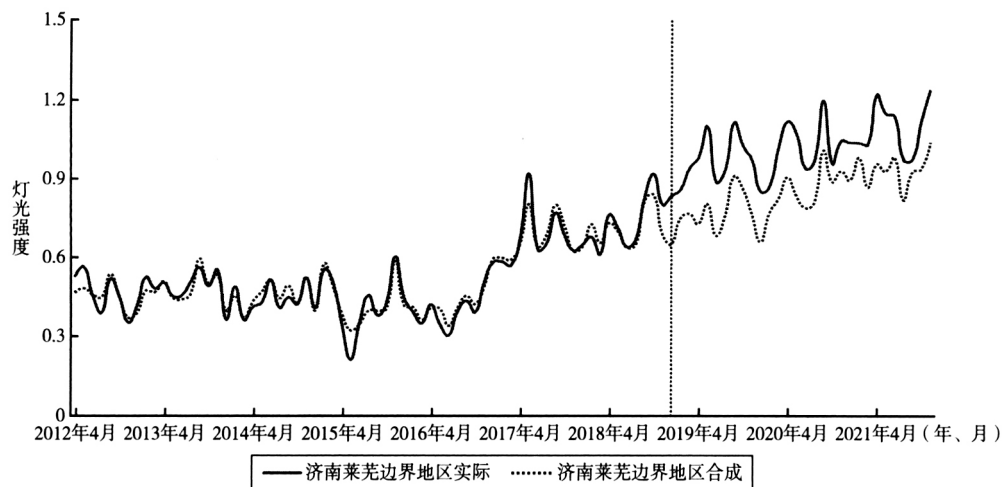


图 4 更换控制组为全国边界地区

最后,在机器学习算法稳健性方面,我们也进行了检验。具体而言,我们将合成控制中使用的机器学习算法由弹性网络算法更换为 Lasso 算法,以检验本文的合成结果对不同的机器学习算法是否具有稳健性。Lasso 算法是一种压缩估计,如同弹性网络算法, Lasso 算法也在目标函数上附加一个惩罚函数,但仅包含 L1 范式的惩罚项。正因为目标函数中 L1 惩罚项的存在,使得 Lasso 回归会压缩回归系数,一些回归系数会为零,其本质也是通过牺牲模型的无偏性来换取更好的预测精度。Lasso 回归的超参数调参过程与弹性网络模型保持一致,在获得 Lasso 模型最优超参数后,便可以利用该参数拟合惩罚回归模型。

使用 Lasso 算法进行合成控制的结果如图 5 所示。可以看出,在更换了机器学习算法后,结论依然保持稳定,处理后政策效果年均均为 15.63%。不同的稳健性检验结果证实了,在更换边界地区度量方式、对照组选取区域以及机器学习算法后,本文结论具有很好的稳健性。由此证明,济南和莱芜合并这一外生政策冲击,对济南莱芜边界地区的经济增长具有显著促进作用。

(三) 安慰剂检验

在得到济南莱芜边界地区实际值和合成值的差值后,很自然存在这样一种猜想:如果随机选择一个边界地区作为处理组,并用其他边界地区作为控制组进行合成控制,那么是否可以得到相同的政策效果?可以预见的是,如果随机选择边界地区进行合成控制的政策效果不明显,而只有济南莱芜边界地区政策效果明显,说明济南莱芜边界地区在政策发生后的确产生了显著的影响。上述安慰剂检验方法,在合成控制法文献中又被称为置换检验方法(Permutation Test),常被用于检验根据合成控制法得到的政策效果是否在统计上显著(Abadie 等,

2010;Clay 等,2020;王贤彬和聂海峰,2010;连玉君和李鑫,2022)。

该方法的核心思想是,随机选取一个控制组中的单元,假设其在 2019 年 1 月也遭遇城市合并的外生冲击,针对该单元同样利用弹性网络模型构造它的合成控制路径,以观测该单元虚假发生的政策效果。在得到控制组中每个单元在虚假的城市合并所带来的政策效果后,将这些控制组单元发生的虚假政策效果与济南莱芜边界地区在真实的城市合并中发生的政策效果进行对比,如果济南莱芜真实城市合并边界地区政策效果显著,而随机选择的虚假城市合并单元政策效果不明显,则可以说明济南莱芜边界地区在政策实施后受到了显著的影响。此外,参考一系列文献的做法(Abadie 等,2010;刘乃全和吴友,2017;王贤彬和聂海峰,2010),在进行安慰剂检验的过程中,如果控制组中某个城市对边界地区的 RMSPE(实际灯光亮度和合成灯光亮度差距平方和均值的平方根)高于济南市—莱芜市 RMSPE 的 2 倍,则可以认为该边界地区在实际政策冲击前没有得到很好的合成,因此该边界地区在 2019 年 1 月后的对照作用较差,故而应当在可视化呈现时剔除这些没有较好合成的地区,仅保留政策冲击前 RMSPE 小于 2 倍济南莱芜边界地区的安慰剂结果。

图 6 展示了安慰剂检验的结果,实线为济南莱芜边界地区在不同时期的实际观测数值减去他们的合成控制反事实路径,而每条虚线则反映了山东省其他每个边界地区在不同时期的实际观测数值减去该边界地区的相应控制组的合成反事实数值。对比实线和虚线可以发现,在 2019 年 1 月两城合并后,济南莱芜边界地区的政策效果(实际观测值减去合成值)显著位于其他边界地区的上方。这一安慰剂检验结果表明,上文估计得到的济南莱芜合并对其边界地区产生了政策效果的结论是可信的。

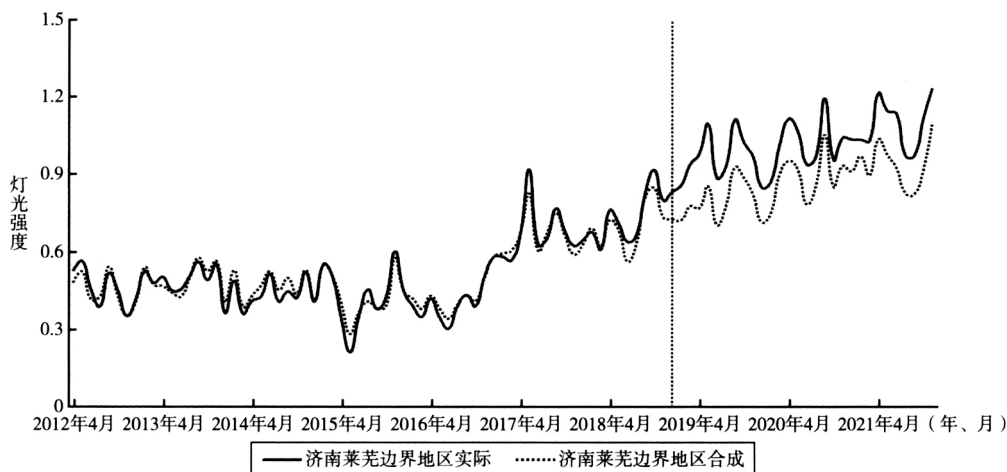


图 5 Lasso 算法下的政策效应

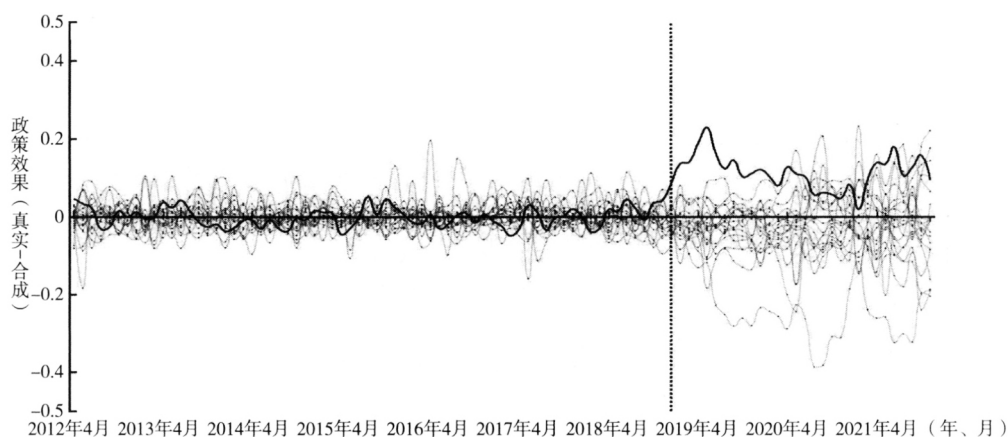


图6 安慰剂检验

在图6中可以看到有个别地区在个别月份的“政策效果”反而超过了济南莱芜边界地区,进一步观察可以看出这是由于这些地区在政策发生前的合成效果较差,因此在合成后的政策效果就存在较大的波动。因此,为了降低干预后噪音成分的影响,参考连玉君和李鑫(2022)最新推荐的做法,针对政策冲击后的实际观测值减去合成值的差值,以政策冲击前的RMSPE为标准进行调整。具体调整方法如下:如果某个城市对边界地区在合成前的RMSPE较大,则证明该边界地区的合成效果不理想,在政策冲击后要对这部分合成不理想的个体施加严格的惩罚,而对合成较好的单元降低惩罚力度,以期实现降低干预后噪音成分对显著性水平的扭曲。而对不同个体施加不同惩罚的具体含义是将政策冲击后实际观测值减去合成值的差值乘以某个系数,这个系数越大代表惩罚越小,反之则代表惩罚越大,主要体现在惩罚因子上。 j 代表第 j 个城市对边界地区, $RMSPE_j^{pre}$ 为第 j 个边界地区在政策冲击前的均方预测误差,针对 $RMSPE_j^{pre}$ 取倒数则得到定义的惩罚因子,该因子反映了对不同单元的惩罚程度,假设一个单元在政策冲击前合成效果好,则 $RMSPE_j^{pre}$ 较小,

代表惩罚力度较小,反之惩罚力度大。

在加入该惩罚因子进行调整后,结果如图7所示。济南莱芜边界地区在城市合并后受到的政策效果,相对于其他单元在伪政策冲击后的虚拟政策效果表现更加显著。该结果表明,随机选取控制组中的一个边界地区,能够得到与济南莱芜边界地区一样的处理效应,是一个小概率事件。而这就反过来验证在济南莱芜合并这一外生政策冲击下,济南莱芜边界地区的经济增长受到了显著正向影响这一结论是可信的。

五、进一步分析

(一)边界两侧异质性政策效应分析

如上文分析中所阐述的,从各方面来讲,济南和莱芜的城市规模差异都很大,这是一次典型的大城市与中小城市之间的合并。为了考察这种大小城市之间的合并对原边界地区两侧是否产生了异质性的影响,在本部分,我们以济南莱芜边界地区的济南边界一侧(垛庄镇、文祖镇,现文祖街道、官庄镇,现官庄街道)和莱芜边界一侧(大王庄镇、雪野镇、茶业口镇)分别作为研究对象进行考察。合成方法仍然使用前文基准模型中的弹性网络算法,超参数调优过程

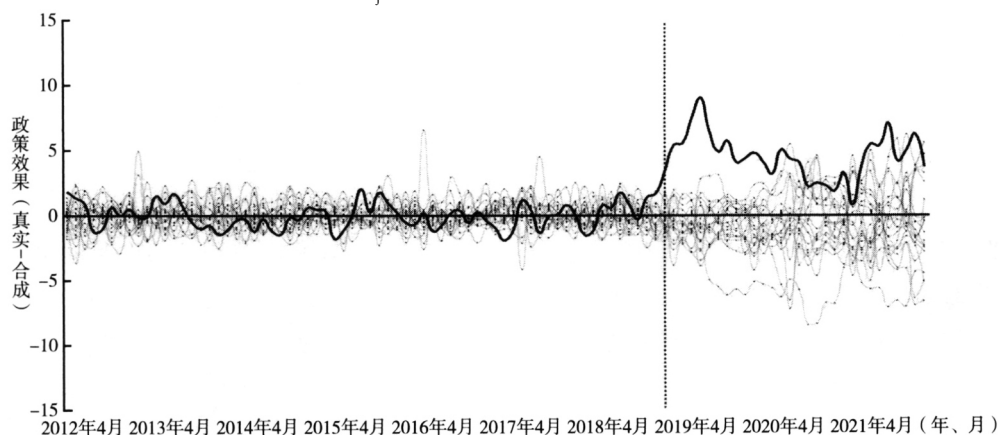


图7 惩罚因子调整后的安慰剂检验^⑦

及合成方法也同上文保持一致。图8(a)展示了济南边界一侧的三个镇作为处理组,山东省内其他边界地区作为控制组的合成结果。从中可以看出济南边界一侧三镇在2019年1月受到合并政策冲击后,其灯光亮度的实际观测值(实线)与合成值(点线)之间存在显著的正向差异,说明济南莱芜的合并确实促进了边界地区济南一侧的经济增长。而图8(b)则展示了莱芜边界一侧三镇作为处理组,山东省内其他边界地区作为控制组的合成结果。从中可以看出,莱芜边界三镇夜间灯光亮度实际值(实线)与合成值(点线)相比并没有出现明显的变化,即济南莱芜的合并对边界地区莱芜一侧经济增长的促进作用并不明显。

对于济南与莱芜的合并对原边界地区两侧产生的异质性政策效果的原因,本文从济南和莱芜原本的城市规模和发展模式进行分析。在中国,地方政府之间存在围绕经济增长的竞争(周黎安,2007),而且这种竞争还是依赖于地区级相对排序的“锦标赛”模式,从而导致边界地区发展相对受到忽视,边界地区发展落后于中心城区(周黎安和陶婧,2011;郭峰等,2023b)。但这是一般意义上的边界地区发展落后的体制逻辑,当相邻两个城市发展规模和阶段存在较大差异时,这一逻辑会有一些新的变化。对济南市而言,为了避免其大城市经济规模集聚和经济增长的红利溢出到其他城市,济南市政府没有激励在其靠近莱芜市的边界地区投入更多资源,从而导致这一边界地区发展水平相对落后。而对莱芜市而言,其城市规模体量很小,市场有限,为了更好地发展本地经济,就需要主动靠近与其距离较近的大城市济南市,承接济南市经济集聚的红利辐射。但同时,莱芜市也要避免经济资源被济南市虹吸,从而就会在靠近济南市,但又属于莱芜市的边界地区优先投入更多资源,从而导致两城市边界地区莱芜一侧发展水平相对更高。而济南市和莱芜市合并之后,原边界地区济南一侧不再是边界地区,行政分割效应消除,迎来发展机遇,经济快速增长。而对于原边

界地区莱芜一侧则情况更为复杂一些,虽然两城市合并后,其接受济南市主城区经济集聚辐射的可能性变高,但也存在不利的因素:一方面,该地区本来就是莱芜市优先发展方向,行政边界消失后,这一地区优先发展的意义下降;另一方面,少了行政壁垒的保护,该地区的经济资源还有可能被虹吸到济南主城区。总体而言,没有观察到济南莱芜合并对原边界地区莱芜一侧的积极影响,就是完全符合上述逻辑的结果^③。

从一般意义上讲,上述结果隐含的逻辑是大城市靠近小城市的边界地区,经济增长水平反而相对更差,而小城市靠近大城市的边界地区,发展水平反而相对更好。为了进一步验证这一逻辑,我们以2015年的卫星灯光为例,构造全国范围内任意两个接壤城市的边界地区经济增长水平对比。图9绘制了城市规模对比和边界经济发展水平对比的拟合图^④。其中,城市规模参考Guo等(2022)、郭峰等(2023b)等文献的做法,将城市灯光亮度达到一定阈值的地区的灯光亮度汇总,这样处理可以提取出真正的城市化区域的经济规模,而不是整个城市管辖行政区的经济规模,这两者在中国语境下有着非常明显的区别。而且考虑到不同城市之间的经济发展水平可能存在系统性的差异,在比较大小城市边界两侧经济发展水平时,我们以各自城市非边界地区的灯光亮度进行了“校准”,即大小城市边界亮度之比= $(\text{大城市边界镇灯光平均亮度}/\text{大城市非边界镇灯光平均亮度})/(\text{小城市边界镇灯光平均亮度}/\text{小城市非边界镇灯光平均亮度})$ 。图9中高度负向的拟合关系,意味着大小城市边界地区的相对发展水平差异与他们之间的城市规模差异确实是反向的关系。通俗来说,对于大城市与小城市交界的地区,大城市往往更缺乏动力去发展。这充分印证了上文关于大小城市地方政府对它们之间的边界地区重视程度存在差异的分析思路。而同时,这一结果又反过来证明前文关于济南莱芜合并对边界两侧产生异质性影响是完全可以理解且置信的。

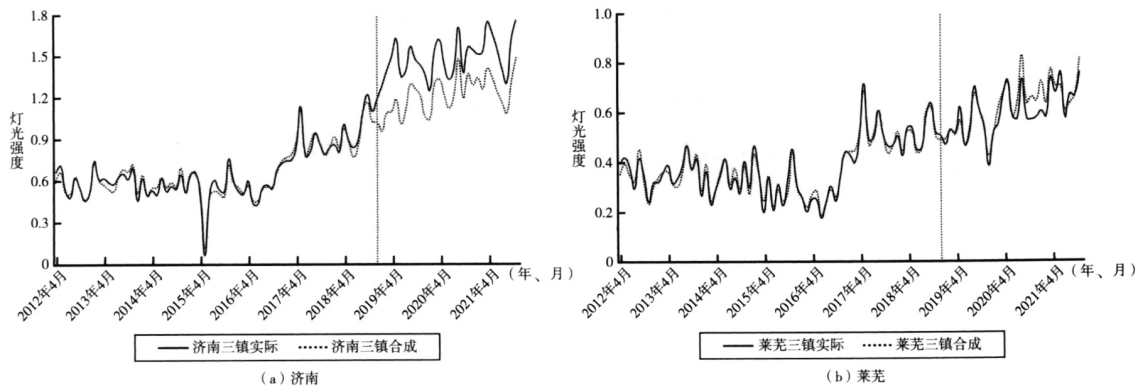


图8 济南莱芜边界两侧灯光分别合成

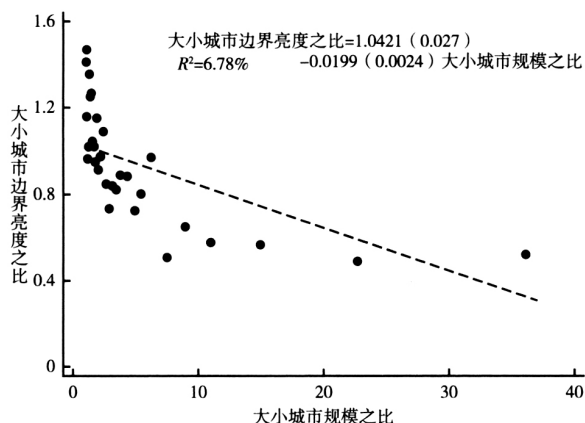


图9 2015年全国城市对规模对比
与边界地区相对发展水平对比

(二) 中心城区异质性政策效应分析

考察完全合并政策对济南莱芜边界两侧分别带来的影响后,很自然的疑问便是该合并政策对济南、莱芜的中心城区会带来怎样的影响。通过对比合并对边界地区和中心地区的影响,也能对本文的核心逻辑进行进一步的论证。为此,本文参考郭峰等(2023b)定义中心城区的方式,分别以扣除边界镇之外的乡镇作为济南、莱芜的中心城区的处理组,以除济南、莱芜外全国其他地级市中心城区作为控制组进行合成,进而考察济南莱芜的合并对济南和莱芜两个城市的中心城区产生的影响。此时的结果如图10所示,从中可以看出合并政策的发生,对于济南中心城区也产生了一些正向效果(5.51%),远高于对于莱芜中心城区的效果(0.25%)。由此说明,济南莱芜合并的发生,无论是对中心城区的影响还是对边界地区的影响方面,在济南市都有着更大的正向政策效果,而对莱芜市政策效果则不明显。由于大城市拥有更高的人口密度和经济活动密度,能够实现更高的生产效率和资源利用效率,从而吸引了中小城市的人口和产业向大城市迁移。在合并之前,莱芜和济南作为两个城市,还存在行政壁垒作为护城河,不至于太多资源从莱芜市被虹

吸到济南市。而在济南和莱芜合并之后,行政壁垒被打破,户籍制度等约束条件弱化或消失,市内通勤成本大幅下降,要素资源流动不再受行政地理边界的限制。这使得小城市的居民迁移成本变得更低,能够更便捷地通过搬迁享受到更丰富的教育、医疗和文化资源,更大的就业市场和更具多样化的消费。综上所述,大城市对中小城市的中心城区也形成了一定的虹吸效应。从而济南中心城区在济南莱芜的合并中,相对收益更多,而莱芜市则并无获益。而对比中心城区和边界地区受到的不同影响,可以看出边界地区受到的正向影响高于中心城区,这进一步证明了本文关于行政合并对边界地区影响的核心逻辑。

(三) 外部有效性检验

1. 样本外测试——预留一年数据进行测试

在基准模型中,我们重点关注在2019年1月济南莱芜合并政策公布之后,济南莱芜边界地区观测值与合成值之间是否存在差异。但是存在和传统合成控制方法一样的问题,即这部分差异既可能是由于真实的政策效果造成的,也可能是模型“过拟合”所导致的。尤其是在本文的应用情境中,由于缺少协变量,因过拟合带来的“伪政策效果”的可能性会进一步加大。这里的过拟合具体是指,弹性网络算法在使用2012年4月~2018年12月数据进行训练时,过度学习了这部分数据的特征,即过度拟合或复制了政策干预前的结果变量,从而可能使得该模型反而无法很好地拟合2019年1月以后的数据。最终导致即便没有合并的发生,模型拟合出的2019年1月以后的反事实路径与济南莱芜边界地区所具有的实际路径也会存在较大差异。

通俗来说,这里出现的是由于过拟合带来的无法泛化的问题,在经济学中也称为“外部有效性”问题(郭峰和陶旭辉,2023)。言下之意是,模型虽然在济南莱芜合并政策前的样本中具有很好的拟合能力(满足内部有效性),但是在该样本之外却可能因为缺乏预测力导致结果产生偏差(外部有效性不足)。

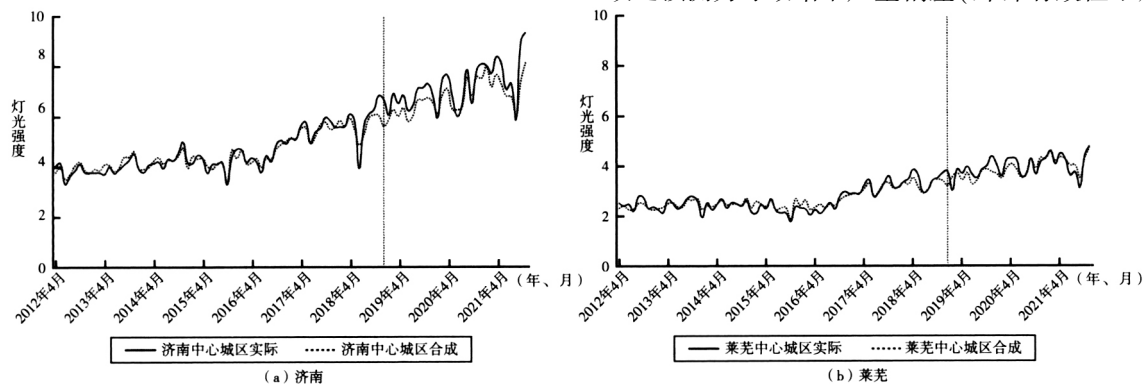


图10 济南和莱芜中心城区灯光分别合成

过拟合现象一旦发生,则意味着处理组和合成控制组在干预后的差异可能并不全部由政策所导致,算法本身的过拟合现象也对差异部分存在一定的贡献,从而对实际政策识别效果起到混淆和干扰的影响。

为解决该问题,我们更换了数据集合的切分方式,将原始训练集中实际政策发生前一年的样本切分出来作为模型外测试集。即仅使用2012年4月~2017年12月的样本作为训练集,构建济南莱芜边界地区与其他城市对边界地区之间的合成关系,而将2018年1月~2018年12月样本作为模型外测试集。通过观察济南莱芜边界地区观测值和合成值在这一时间段上的表现,便可以识别出在干预时间后该地区灯光亮度实际值和合成控制组之间的差异,究竟是训练集上过拟合因素所导致,还是由于真实政策影响所导致。若济南莱芜边界地区观测值与合成值在测试集上贴合紧密,则能很好地证明过拟合现象并不严重,从而证明2019年1月后济南莱芜边界地区的观测值与合成值的差值确实是由实际政策所带来的,进而避免“伪政策效果”问题。这种对数据集进行划分,预留一部分样本对因果识别结果进行测试评估,也是机器学习对包括合成控制法在内的因果识别实践的一个重要启示(郭峰和陶旭辉,2023)。

图11展示了重新切分数据集后的结果,本文利用训练集中2012年4月~2017年12月的灯光亮度作为训练集进行模型训练,进而观察模型在2018年1月~2018年12月的测试集上的表现。从图11中可以看出,总体而言,在2018年1月~2018年12月的测试集中,济南莱芜边界地区观测值和合成值并无显著差异,说明过拟合因素对本研究实证结果影响甚微。而这一结果更进一步证明在2019年1月~2021年12月数据中,济南莱芜边界地区实际观测值

和合成值之间的差异,就可被解释为济南莱芜合并的政策效果。

在图11中,我们也可以看出,将2018年作为测试集时,2018年的最后几个月,其实济南莱芜边界地区的实际值有略微高于其反事实合成值的趋势,这里很可能是因为该政策会被提前预期到。实际上,在中央政府正式批准济南莱芜的合并之前,地方政府已经进行了一些准备工作,提前发力。不过,还有一种可能性是我们的反事实估计存在过拟合,如果过拟合存在,也完全可能会发生这种现象。因此为了对此进一步进行检验,我们将验证集分别设置为2017年和2016年来进行考察。首先,如图12(a)中虚线所框选部分所示,我们采用2017年1月~2018年1月作为评估过拟合的区间,即使用2012年4月~2016年12月和2018年2月~2018年12月作为训练集进行模型训练,观察模型在2017年1月~2017年12月的测试集上的表现。从图12(a)中可以看出,2017年的实际观测值和合成值并无显著差异,因此在排除政策被提前预期的前提下,更加干净地识别了机器学习算法的过拟合对本文结论影响甚微。图12(b)与(a)逻辑一致,我们将验证集向前推移两年,到2016年1月~2017年1月,结论依旧稳健^⑩。图11和图12的结果也进一步证明,济南莱芜的合并确实被提前预期到了,因此在政策正式落地前,政策效果已经有所显现。同时,政策提前被预期到并不会影响本文基准结果的成立,反而会使得本文的结果低估了真实的处理效应,即政策实际产生的处理效应应当比论文所汇报出的实证结果还要高。

2. 更多案例:长春—公主岭案例

事实上,关于大小城市的合并,虽然发生的频率不高,但济南与莱芜的合并也并非孤例。为了进一步检验本文的研究结论和研究方法的外部有效性,我

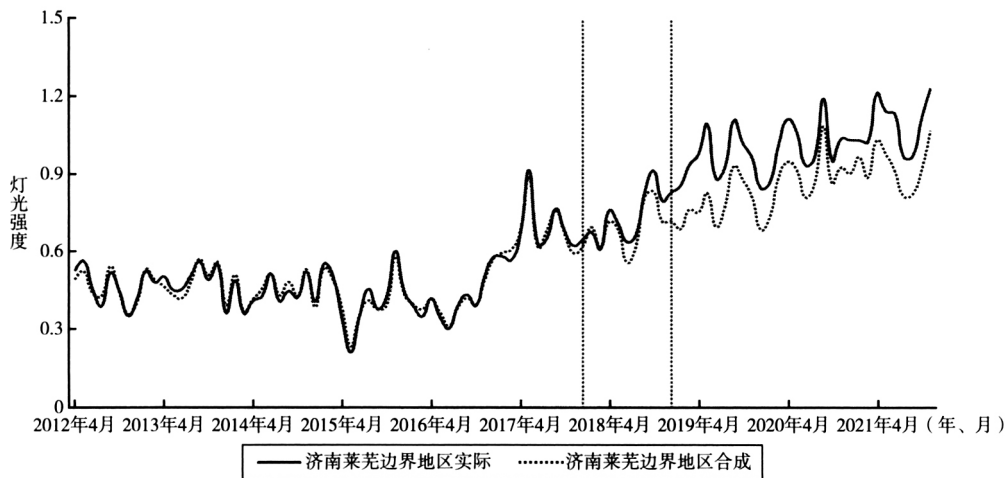


图11 样本外测试:2018年为测试集

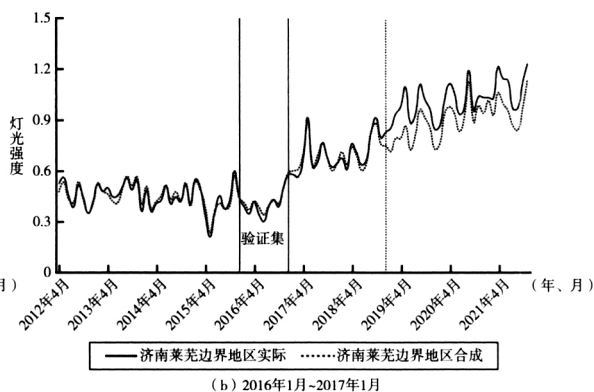
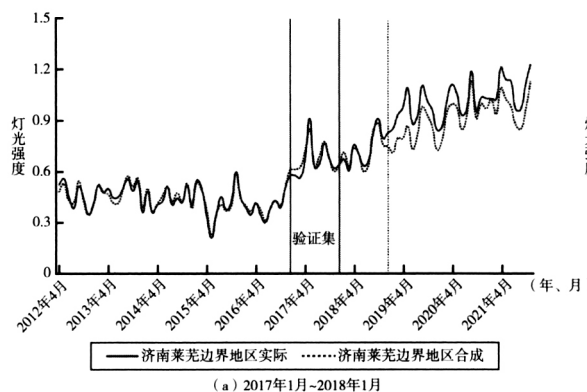


图 12 样本外测试:更换验证集

们将长春代管公主岭作为另外一个案例进行考察。具体而言,2020年6月19日,经国务院批复同意,将原由四平市代管的县级公主岭市改由长春市代管。因此本文分别以长春公主岭边界地区长春一侧边界地区、公主岭一侧边界地区分别作为处理组,而控制组则选取东北三省(辽宁省、吉林省、黑龙江省)其他两两城市边界地区,共计79个^⑩。由于公主岭市是由四平市划分出,则在合成控制的控制组中,排除了长春市和四平市的边界地区,因此实际合成控制中用到的控制组数量为78个。结果如图13所示,边界地区长春一侧实际结果与合成反事实结果差异为年均27.78%,而公主岭一侧的政策效果年均19.22%。从该案例的结果可以看出,城市合并(行政区划调整)虽然对小城市一侧也产生了积极影响,但更大的积极影响仍然主要体现在大城市一侧,这一结论与本文的核心结论大体保持一致^⑪。

六、结论与政策启示

城市是现代经济体系的骨架和非农产业的重要载体,也是经济增长的重要来源,但城市规模的扩张,除受到经济和地理等因素的影响外,还受制于地区行政区划的影响。为了突破行政区划对城市规模扩张的约束,近年来中国的城市行政区划调整频繁,既包括城市辖区内部的区划调整,也包括城市间的

行政区划调整,例如两个城市直接合并。两个城市合并改变了他们之间的原行政边界地区的发展条件,从而可能对边界地区的经济增长产生积极影响。本文以“莱芜并入济南”这一外生冲击事件作为自然实验,利用细颗粒度的卫星灯光数据更精细地刻画城市行政边界地区经济增长水平,并使用机器学习方法(弹性网络算法、Lasso算法)拓展传统合成控制在进行政策评估中的适用边界,考察了济南莱芜合并对他们之间的原边界地区经济增长的影响。

实证研究发现,济南和莱芜的合并确实会对原边界地区经济增长产生正向影响,在更换边界地区度量方式、对照组选取方式以及机器学习算法后,结论始终保持稳健。值得注意的是,本文发现这两个规模差异很大的城市合并,对原边界地区两侧的影响存在异质性效应:合并的正向效应主要体现在大城市济南市边界一侧,而被合并的莱芜市边界一侧则受益甚微。这意味着对核心城市济南市而言,城市合并会打破行政壁垒,拓展了城市规模扩张的行政区划限制,提高了资源空间流动效率,从而促进边界地区经济增长。但对于既依赖济南市城市经济红利辐射,又依靠行政壁垒作为护城河,保证经济资源不被济南市完全虹吸的中小城市莱芜市而言,在经历城市合并后,靠近济南市的原莱芜边界地区,反而

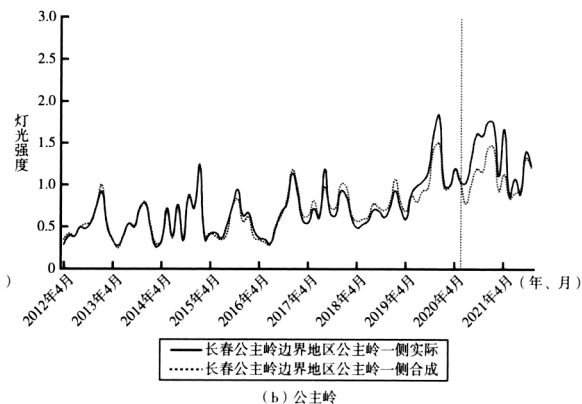
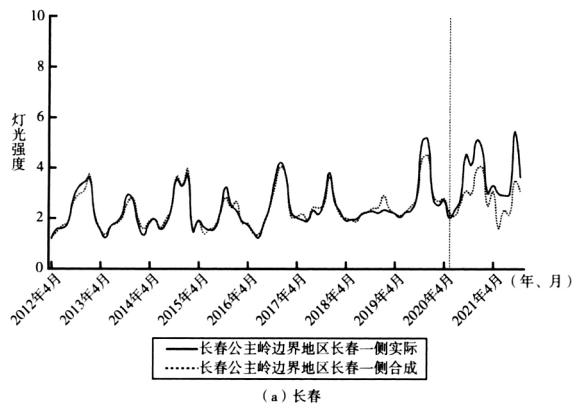


图 13 长春和公主岭边界地区灯光分别合成

没有享受到城市行政区划边界消失的红利。在进一步分析中,本文发现中心城区结论与边界地区结论相似,合并政策同样是对大城市的中心城区有更高的正向效果。此外,为了检验结论的外部有效性,本文选取长春代管公主岭作为新的合并案例进行考察,其结论与本文核心结论也是一致的。

基于合成控制实证结果,本文可以得到以下几点政策启示。第一,研究发现大小城市合并有助于促进城市行政边界地区的经济增长,边界地区的这种增长效应主要来自行政壁垒的消除。本文认为,打破行政壁垒确实可以起到优化资源配置和提高空间治理效率的作用,促进原本处于劣势的城市行政边界地区更快发展。因此,加强大小城市之间的合并有助于削弱上述行政壁垒。一方面,鼓励中央政府加强地方政府间的统筹作用,以自上而下的方式,形成城市间的兼并或合并局面,进而缓解城市间行政壁垒的形成;另一方面,由大小城市政府自发地建立协调机制,以自下而上的方式,实现边界壁垒的消除,促进资源合理配置,进而实现城市行政边界地区的经济发展。

第二,根据本文的分析逻辑,大小城市的合并政策会对大城市和被合并的中小城市造成异质性影响,进而不利于实现区域协调发展和共同富裕。具体来说,大小城市合并政策的边界增长效应主要体现在大城市的一侧,而对中小城市的行政边界地区则没有效果。因此,在制定相应的区域发展政策时,对不同类型的城市不能单纯地实施“一刀切”政策,从而导致更大的区域发展差距形成。应当通盘考虑各城市的经济增长水平和资源禀赋情况,因地制宜地制定相应的促进区域协调发展的政策。对大城市而言,由于城市合并后大城市的边界地区发展相对较快,应及时对大城市的边界地区实施相应的产业发展配套政策。对中小城市而言,由于城市合并后边界地区的发展相对较慢,因此应该通过鼓励投资、吸引人才流入等措施进行更多的要素资源配置,避免造成小城市的产业结构单一,进而促进边界地区的经济增长和人民福祉提升。

第三,本文研究结果的一个隐含结论是,解决城市行政边界地区发展落后的洼地效应,关键在于大城市。小城市有激励在靠近大城市的边界地区投入更多经济资源,而大城市则没有激励发展靠近小城市的边界地区。因此,有必要发挥上一级政府的调节引导作用,促使大城市摒弃“以邻为壑”思路,打开“断头路”,深化区域合作,促进大中小城市经济协同发展。具体来说,中央政府应鼓励大城市加大对行政边界地区的公共物品投资,打破大城市与中小城市之间的行政壁垒,实现区域一体化的发展格局。同时,地方政府在边界地区进行公共投资时,还应紧

盯民生需求,改善人民群众的经济状况,增强地区公共服务供给,推动基本公共服务均等化实现,为资本、劳动力等生产要素在本地落地生根提供强有力的环境保障。

(感谢匿名审稿专家的宝贵意见,文责自负。)

注释:

①数据来源于《中国统计年鉴》和《2022 年中国城市建设状况公报》。

②共涉及 6 个乡镇,原济南市和原莱芜市各 3 个。需要补充说明的是,济南莱芜的合并对边界镇而言并不涉及级别的改变,例如原来“莱芜市—莱城区—大王庄镇”行政级别是正科级,在合并后,该镇变为了“济南市—莱芜区—大王庄镇”,仍是正科级级别。

③弹性网络算法是岭回归算法和 Lasso 回归算法的一种组合,后两者都是弹性网络算法的一种特殊形式,因此这里仅介绍基于弹性网络算法进行合成控制的逻辑和实操要点。

④如上文所述,济南莱芜合并的批复时间是 2018 年 12 月底,2019 年 1 月合并后的地方政府正式挂牌,因此我们将 2019 年 1 月作为政策起始月份,下文同此表述。

⑤处理效应计算方式:首先计算出政策发生后逐月的处理效应=(济南莱芜边界地区实际灯光-济南莱芜边界地区合成控制的反事实灯光)/(济南莱芜边界地区合成控制的反事实灯光);其次,将计算出政策发生后逐月的处理效应在月份层面取平均,得到平均处理效应。

⑥本文附录详见《数量经济技术经济研究》杂志网站,下同。

⑦删除 RMSPE 高于济南莱芜边界地区两倍的其他边界地区。

⑧为了提供更多证据,我们整理了历年的道路矢量数据,并将其和济南莱芜边界镇矢量数据合并,进一步比较了济南边界一侧道路长度相比莱芜边界一侧道路长度在合并前后的变化。虽然采用年度数据,但从中还是可以粗略看出济南一侧边界地区道路在合并后确实显著增加,而莱芜一侧边界地区道路的增加幅度则相对有限。由于篇幅限制,相关证据备索。

⑨另外,在分析城市规模对比和边界经济发展水平对比关系中,为使得不会因为样本太多而影响图形对二者关系的直观反映,图 9 的绘制采用分仓散点图的方式。

⑩实际上,本文还随机选取了其他期间进行样本外测试,结论均一致,限于文章篇幅没有呈现。

⑪由于夏季北半球存在极昼的现象,中国东北地区每年的 5~7 月都会存在失真情况,故此部分剔除了 5~7 月的数据。

⑫由于东北地区数据特殊,导致灯光数据波动较大,因此在图上看起来不是很明显,但实际上,我们计算出政策发生后逐月的处理效应=(长春公主岭边界地区实际灯光-长春公主岭边界地区合成控制的反事实灯光)/(长春公主岭边界地区合成控制的反事实灯光),然后将计算出政策发生后逐月的处理效应在月份层面取平均,得到平均处理效应是 27.78% 和 19.22%,政策效果是相对大且明显的。

参考文献:

[1]付婷婷,张同斌.城市边界扩张能否形成经济增长新动力——以部分省会城市的扩张为例[J].经济科学,2022,(4):50~63.

[2]郭峰,曹友斌,熊云军,吕斌.国家级新区设立与企业空间布局:基于镇级面板数据的分析[J].经济研究,2023a,58(8):191~208.

[3]郭峰,胡军.地区金融扩张的竞争效应和溢出效应——基于空间面板模型的分析[J].经济学报,2016,3(2):1~20.

[4]郭峰,陶旭辉.机器学习与社会科学中的因果关系:一个文献综述[J].经济学(季刊),2023,23(1):1~17.

[5]郭峰,熊云军,石庆玲,王靖一.数字经济与行政边界地区经济发展再考察——来自卫星灯光数据的证据[J].管理世界,2023b,39(4):16~33.

[6]李宝礼,邵师.行政资源重配、区域市场整合与城市经济增长——来自巢湖市并入合肥市的经验证据[J].江西财经大学学报,2019,(5):22~33.

[7]连玉君,李鑫.合成控制法中的安慰剂检验改进研究——基于标准化转换的统计推断[J].统计研究,2022,39(8):115~128.

[8]刘甲炎,范子英.中国房产税试点的效果评估:基于合成控制法的研究[J].世界经济,2013,36(11):117~135.

[9]刘乃全,吴友.长三角扩容能促进区域经济共同增长吗[J].中国工业经济,2017,(6):79~97.

[10]刘修岩,李松林,秦蒙.开发时滞、市场不确定性与城市蔓延[J].经济研究,2016,51(8):159~171+186.

[11]卢盛峰,陈思霞.政府偏袒缓解了企业融资约束吗?——来自中国的准自然实验[J].管理世界,2017,(5):51~65+187~188.

[12]唐为.分权、外部性与边界效应[J].经济研究,2019,54(3):103~118.

[13]唐为,王媛.行政区划调整与人口城市化:来自撤县设区的经验证据[J].经济研究,2015,50(9):72~85.

[14]王文凯,任元明.边界效应、市场潜力与经济发展差异:基于重庆市直辖的研究[J].财贸经济,2022,43(3):128~142.

[15]王贤彬,聂海峰.行政区划调整与经济增长[J].管理世界,2010,(4):42~53.

[16]魏守华,杨阳,陈珑隆.城市等级、人口增长差异与城镇体系演变[J].中国工业经济,2020,(7):5~23.

[17]魏守华,周山人,千慧雄.中国城市规模偏差研究[J].中国工业经济,2015,(4):5~17.

[18]许政,陈钊,陆铭.中国城市体系的“中心-外围模式”[J].世界经济,2010,33(7):144~160.

[19]尹恒,徐琰超.地市级地区间基本建设公共支出的相互影响[J].经济研究,2011,46(7):55~64.

[20]余华义,侯玉娟,洪永森.城市辖区合并的区域一体化效应——来自房地产微观数据和城市辖区经济数据的证据[J].中国工业经济,2021,(4):119~137.

[21]张楠,高明,杨琳.突破传统行政区划与经济发展——来自城市区界重组的证据[J].数量经济技术经济研究,2023,40(11):158~179.

[22]周黎安.中国地方官员的晋升锦标赛模式研究[J].

经济研究,2007,(7):36~50.

[23]周黎安,陶婧.官员晋升竞争与边界效应:以省区交界地带的经济发展为例[J].金融研究,2011,(3):15~26.

[24]Abadie A., Diamond A., Hainmueller J., 2010, Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program[J], Journal of the American Statistical Association, 105(490), 493~505.

[25]Abadie A., Gardeazabal J., 2003, The Economic Costs of Conflict: A Case Study of the Basque Country[J], American Economic Review, 93(1), 113~132.

[26]Ahlfeldt G. M., Redding S. J., Sturm D. M., Wolf N., 2015, The Economics of Density: Evidence from the Berlin Wall[J], Econometrica, 83(6), 2127~2189.

[27]Arkhangelsky D., Athey S., Hirshberg D. A., Imbens G. W., Wager S., 2021, Synthetic Difference-in-Differences[J], American Economic Review, 111(12), 4088~4118.

[28]Athey S., Imbens G. W., 2017, The State of Applied Econometrics: Causality and Policy Evaluation[J], Journal of Economic Perspectives, 31(2), 3~32.

[29]Athey S., Imbens G. W., 2019, Machine Learning Methods That Economists Should Know About[J], Annual Review of Economics, 11(1), 685~725.

[30]Blesse S., Baskaran T., 2016, Do Municipal Mergers Reduce Costs? Evidence from a German Federal State[J], Regional Science and Urban Economics, 59, 54~74.

[31]Clay K., Egedesø P. J., Hansen C. W., Jensen P. S., Calkins A., 2020, Controlling Tuberculosis? Evidence from the First Community-wide Health Experiment[J], Journal of Development Economics, 146, 102510.

[32]Cole M. A., Elliott R. J. R., Liu B., 2020, The Impact of the Wuhan Covid-19 Lockdown on Air Pollution and Health: A Machine Learning and Augmented Synthetic Control Approach[J], Environmental and Resource Economics, 76(4), 553~580.

[33]Doudchenko N., Imbens G. W., 2016, Balancing, Regression, Difference-in-Differences and Synthetic Control Methods: A Synthesis[R], NBER Working Paper, No. 22791.

[34]Gibson J., Olivia S., Boe-Gibson G., 2020, Night Lights in Economics: Sources and Uses[J], Journal of Economic Surveys, 34(5), 955~980.

[35]Gibson J., Olivia S., Boe-Gibson G., Li C., 2021, Which Night Lights Data Should We Use in Economics, and Where? [J], Journal of Development Economics, 149, 102602.

[36]Guo F., Shi Q., Wang J., Xiong Y., 2022, Digital Economy and Decentralization of Cities: Evidence from China[R]. Working Paper.

[37]Guo J., Zhang Z., 2019, Does Renaming Promote Economic Development? New Evidence from a City-renaming Reform Experiment in China[J], China Economic Review, 57, 101344.

[38]Henderson J. V., Storeygard A., Weil D. N., 2012, Measuring Economic Growth from Outer Space[J], American Economic Review, 102(2), 994~1028.

[39]Henderson J. V., Wang H. G., 2005, Aspects of the Rural-Urban Transformation of Countries[J], Journal of Economic

Geography, 5(1), 23 ~ 42.

[40] Jia J., Liang X., Ma G., 2021, Political Hierarchy and Regional Economic Development: Evidence from a Spatial Discontinuity in China[J], Journal of Public Economics, 194, 104352.

[41] Kumar A., Liang C. Y., 2019, Credit Constraints and GDP Growth: Evidence from a Natural Experiment[J], Economics Letters, 181, 190 ~ 194.

[42] Lucas R. E., 2004, Life Earnings and Rural-Urban Migration[J], Journal of Political Economy, 112(S1), S29 ~ S59.

[43] Mühlbach N. S., Nielsen M. S., 2021, Tree-based Synthetic Control Methods: Consequences of Moving the US Embassy [J/OL], arXiv:1909.03968.

[44] Reingewertz Y., 2012, Do Municipal Amalgamations Work? Evidence from Municipalities in Israel[J], Journal of Urban Economics, 72(2), 240 ~ 251.

[45] Xu C., 2011, The Fundamental Institutions of China's Reforms and Development[J], Journal of Economic Literature, 49(4), 1076 ~ 1151.

Merger of Large and Small Cities and Economic Growth in Boundary Areas: Evaluation by Synthetic Control Based on Machine Learning Algorithms

Guo Feng Lv Bin Xiong Yunjun Tao Xuhui

Abstract: The slow economic growth in areas located on administrative boundaries due to factors such as the division of administrative power and local protectionism is one of the major obstacles to achieving the great circulation of the national economy and the high-quality development of the regional economy. Adjusting administrative divisions makes it possible to accelerate the economic development of administrative boundary areas. At present, there have been many studies on changing counties into districts, the merger of two districts, and the establishment of development zones, but there has been a lack of detailed research on whether the adjustment of administrative divisions, especially the merger of small cities by large cities, can reduce the poverty of the boundary areas.

Using the Laiwu merger into Ji'nan as a natural experiment, this study investigates the impact of the small cities merged into large cities on the economic growth in the original boundary area using fine-grained satellite lighting data and a synthetic control method based on machine learning. The results reveal that the merger of Ji'nan and Laiwu has had a positive impact on the economic growth of the boundary area—the actual value of the light brightness of the boundary area is on average 10.8% higher than the predicted counterfactual synthetic value. The conclusions remain robust after replacing the boundary area metric, changing the control group, and using the machine learning method.

Further, the study examines the treatment heterogeneity effect on both sides of the boundary located in large and small cities. The results reveal that the positive effects of mergers are mainly on the boundary located in the core city of Ji'nan, while the boundary located in the small city of Laiwu benefits very little from mergers. The likely reason for this is the competition for economic growth between local governments in China, which relies on the tournament model of relative ordering at the regional level. Before the merger, the Ji'nan government lacked incentives to invest more resources in its boundary area near Laiwu, which led to a relatively low level of development in this area. Laiwu, with its small city size, tends to proactively approach Ji'nan to absorb the spillover effects of the economy from Ji'nan, which made boundary areas located in Laiwu more developed than expected. After the city merger, the administrative division effect is eliminated, and the Ji'nan boundary area has opportunities for rapid economic growth.

Based on the results, this study puts forward three policy recommendations. First, as strengthening the merger between large and small cities helps to weaken the administrative barriers mentioned above, the central government should strengthen the coordination among local governments and encourage the governments of large and small cities to establish a coordination mechanism on their own initiative to promote the rational allocation of resources, thereby realizing the economic development of the areas along the administrative boundaries of cities. Second, the government should consider the level of economic growth and resource endowment of each city and formulate corresponding policies to promote coordinated regional development. Third, governments at the higher level must play a role in coordinating to encourage large cities to abandon the "beggar-thy-neighbor" mentality, open up the broken roads, deepen regional cooperation, and promote the synergistic development of the economies of cities of all sizes.

Key words: administrative boundaries; satellite lighting data; synthetic control method; machine learning